Classical capacity of fermionic product channels

Sergey Bravyi[1,*]

[1] *Institute for Quantum Information, California Institute of Technology, Pasadena, 91125 CA, USA.*

(Dated: February 1, 2008)

We study multi-qubit quantum channels that can be represented as a product of one-mode fermionic attenuation channels. An explicit formula for the classical capacity C_1 and for the minimum output entropy S_{min} of these channels is proposed. We compute S_{min} analytically for any number of qubits under assumption that the minimum is achieved on a Gaussian input. Apart from that, a simple numerical method for evaluating S_{min} is developed. The method is applicable to any channels that are sufficiently noisy. For fermionic product channels the proposed formula for S_{min} agrees with the numerical results with a precision about 10^{-9}.

PACS numbers:

I. INTRODUCTION AND SUMMARY OF RESULTS

Transmission of a classical information through a quantum communication channel is one of the basic problems studied in the quantum information theory. Suppose a sender (Alice) wants to send a string of n classical bits through a one-way noisy quantum channel Φ to a receiver (Bob). Suppose Alice is allowed to use the channel only m times. A number R is called an achievable rate if for any $\epsilon, \delta > 0$ there exists a transmission protocol with $n/m > R-\delta$ and an error probability smaller than ϵ. The classical capacity of Φ is defined as the minimum number C such that $R \leq C$ for any achievable rate R. A closely related quantity is a *one-shot capacity* C_1 which is defined analogously to C with a restriction that Alice uses only *product states* for signaling. It has been shown that $C(\Phi) = \lim_{m\to\infty} C_1(\Phi^{\otimes m})/m$ and conjectured that actually $C(\Phi) = C_1(\Phi)$, i.e., C_1 is additive under tensor product of channels. The Holevo-Schumacher-Westmoreland theorem [1, 2] states that

$$C_1(\Phi) = \sup_{\mathcal{E}} \chi(\mathcal{E}) \tag{1}$$

where $\mathcal{E} = \{p_a, \rho_a \in Im(\Phi)\}$ is a probabilistic ensemble of states from the image of Φ (a set of all output states), and $\chi(\mathcal{E})$ is the Holevo quantity

$$\chi(\mathcal{E}) = S\left(\sum_a p_a \rho_a\right) - \sum_a p_a\, S(\rho_a). \tag{2}$$

Important examples of quantum channels for which an explicit formula for the classical capacity is known are product of one-qubit unital channels [3] and products of bosonic attenuation channels [4]. In the latter case a restriction on a power of the input signal has to be imposed to regularize the capacity. An explicit formula for the one-shot capacity is also conjectured for bosonic attenuation channels combined with a classical Gaussian noise [5].

Quantum channels explored in this paper are direct fermionic analogues of bosonic product channels studied in [4]. In contrast to their bosonic counterparts, fermionic modes are described by a finite-dimensional Hilbert space, so one does not need to regularize the capacity.

An algebra of observables of n fermionic modes formally coincides with the one of n qubits and can be conveniently described by generators $\hat{c}_1, \hat{c}_2, \ldots, \hat{c}_{2n}$,

$$\begin{aligned} \hat{c}_{2j-1} &= \sigma_1^z \otimes \cdots \otimes \sigma_{j-1}^z \otimes \sigma_j^x \otimes I \otimes \cdots \otimes I, \\ \hat{c}_{2j} &= \sigma_1^z \otimes \cdots \otimes \sigma_{j-1}^z \otimes \sigma_j^y \otimes I \otimes \cdots \otimes I. \end{aligned} \tag{3}$$

Here σ_j^α are the Pauli operators on the qubit j. This is the well-known Jordan-Wigner transformation that is often used to map a system of spins (qubits) into a system of fermions. The generators $\hat{c}_1, \ldots, \hat{c}_{2n}$ obey the Clifford algebra commutation rules

$$\hat{c}_p\, \hat{c}_q + \hat{c}_q\, \hat{c}_p = 2\, \delta_{pq} I, \quad \hat{c}_p^\dagger = \hat{c}_p.$$

An arbitrary linear operator acting on n qubits can be uniquely represented as a linear combination of 4^n monomials

$$\hat{c}(\mathbf{x}) = \hat{c}_1^{x_1}\, \hat{c}_2^{x_2} \cdots \hat{c}_{2n}^{x_{2n}},$$

where $\mathbf{x} = (x_1, \ldots, x_{2n})$ is a binary string of $2n$ bits.

Define a quantum channel Φ by the following rules

$$\begin{aligned} \Phi(I) &= I, \\ \Phi(\hat{c}_p) &= b_p\, \hat{c}_p, \quad p = 1, \ldots, 2n, \\ \Phi(\hat{c}(\mathbf{x})) &= \prod_{p\,:\,x_p=1} b_p\, \hat{c}(\mathbf{x}). \end{aligned} \tag{4}$$

Here $0 \leq b_1, \ldots, b_{2n} \leq 1$ are $2n$ real parameters specifying the channel (attenuation coefficients). A proof that the linear map defined in Eq. (4) is indeed a quantum channel (i.e. a trace preserving completely positive map) as well as some motivation of this definition will be given

*Electronic address: serg@cs.caltech.edu

in Section II. It suffices to mention now that Φ has a clear product structure. One can say that the system of n qubits is 'partitioned' into $2n$ separated subsystems (Majorana fermionic modes) that are described by the operators $\hat{c}_1, \ldots, \hat{c}_{2n}$. Each subsystem is transmitted independently through its own 'wire' described by a quantum channel $\hat{c}_p \to b_p \hat{c}_p$. However this product structure should not be mixed with the tensor product of quantum channels. Indeed, we shall see in Section II that Krauss operators corresponding to different 'wires' do not commute. Accordingly, one can not expect that the capacity of the channel Eq. (4) will be additive (a capacity of an individual wire is not even a well-defined quantity).

To compute the maximum at Eq. (1) we shall use a certain variational family of states, namely, fermionic Gaussian states [6, 7]. A Gaussian state ρ is completely specified by its first and second moments $\mathrm{Tr}(\rho\, \hat{c}_p)$, $\mathrm{Tr}(\rho\, \hat{c}_p\, \hat{c}_q)$, see Section III for a strict definition. In particular, the von Neumann entropy $S(\rho)$ is a simple function of these moments. We will see that the product fermionic channels map Gaussian states into Gaussian states. Define a *Gaussian capacity* as

$$C_1^g(\Phi) = \sup_{\mathcal{E}_g} \chi(\mathcal{E}_g), \tag{5}$$

where $\mathcal{E}_g = \{p_a, \rho_a \in Im(\Phi)\}$ is an ensemble of Gaussian states (probabilities p_a may be arbitrary though). This is the best transmission rate that can be achieved if Alice uses only product Gaussian states for signalling. Obviously, $C_1^g \leq C_1$. Our main result is an explicit formula for the Gaussian capacity.

Theorem 1. *Let Φ be a fermionic product channel acting on n qubits with attenuation coefficients $0 \leq b_1, \ldots, b_{2n} \leq 1$. Denote $b_j^{\downarrow}$ the j-largest coefficient. Then*

$$C_1^g(\Phi) = n - H\left(\frac{1+b_1^{\downarrow}}{2}\right) - \sum_{j=1}^{n-1} H\left(\frac{1+b_{2j}^{\downarrow}\, b_{2j+1}^{\downarrow}}{2}\right), \tag{6}$$

where $H(x) = -x\log(x) - (1-x)\log(1-x)$ is the Shannon binary entropy. Moreover, $C_1(\Phi) = C_1^g(\Phi)$ if $n \leq 2$.

A proof of the theorem as well as some additional results that are valid for Gaussian input states is given in Section IV.

The theorem implies that the optimal signaling ensemble is Gaussian for $n \leq 2$. Can one achieve a higher transmission rate for $n > 2$ by using non-Gaussian ensembles? To get some intuition about it, we have tried to calculate C_1 numerically for three and four qubits, see Section V. The numerical results strongly suggest that $C_1 = C_1^g$, i.e., that we gain nothing from using non-Gaussian signaling states. It might seem rather surprising because mixed Gaussian states are known to be the maximally noisy states (in terms of their von Neumann entropy) for fixed first and second moments. On the other hand, one can easily see from Eq. (4) that the channel Φ washes out correlations among large number of modes such that l-mode correlators acquire a factor b^l, where b is a typical value of b_p. It means that two-mode correlations is the best place to keep the information that has to be sent through the channel. In this respect Gaussian states are very promising candidates, because they have *only* two-mode correlations. Therefore, there might be a subtle tradeoff between large entropy of Gaussian states and their special correlation structure that makes them optimal for signaling.

In general, a computation of C_1 (even numerical) is an extremely hard problem. Fortunately, the fermionic product channels possess a special symmetry property known as *covariance*, see [8, 9]. For covariant n-qubit channels the one-shot capacity is given by

$$C_1(\Phi) = n - S_{min}(\Phi), \tag{7}$$

where

$$S_{min}(\Phi) = \min_{\rho} S(\Phi(\rho)) \tag{8}$$

is the minimum output entropy. Our derivation of the Gaussian capacity is actually a computation of the minimum output entropy for Gaussian input states. The analogous problem has been already solved for bosonic Gaussian channels [10] and a conjecture has been made that the global minimum of the output entropy is achieved on a Gaussian input. Whether or not this conjecture is true in the fermionic world is an open question (which is equivalent to an equality $C_1 = C_1^g$).

In Section V we describe a simple algorithm that allows one to evaluate $S_{min}(\Phi)$ for many quantum channels. Applicability of the algorithm is not restricted to the fermionic product channels. In general, the algorithm works well if Φ is not too close to an ideal channel. The main idea is to exploit positivity of the relative entropy and iteratively minimize $S(\Phi(|\psi\rangle\langle\psi|))$. More specifically, we construct a sequence of pure states $|\psi_0\rangle, |\psi_1\rangle, |\psi_2\rangle, \ldots$, such that $|\psi_{j+1}\rangle$ is the largest eigenvector of an operator $\Phi^*(\log \Phi(|\psi_j\rangle\langle\psi_j|))$. The initial state $|\psi_1\rangle$ is chosen randomly. One can easily show that $S(\Phi(\psi_{j+1})) \leq S(\Phi(\psi_j))$, i.e., the corresponding sequence of output entropies is non-increasing. One should repeat the iterations until the output entropy keeps decreasing.

In contrast to local search methods, such as the gradient descent, the iterative minimization is able to escape from local minimums of the objective function. However, one can not guarantee that the iterations converge to the global minimum, see Section V for more detailed discussion.

In Section V.B we use the algorithm to find the minimum output entropy of the fermionic product channels for three-qubit and four-qubit systems. The corresponding value of the capacity C_1 matches the Gaussian capacity C_1^g with a precision about 10^{-9} which has the same order of magnitude as the numerical noise. It is therefore fair to make a conjecture that $C_1 = C_1^g$ in general.

The rest of the paper is organized as follows. Section II contains all the necessary facts about fermionic product

channels. In Section III we review some basic properties of Gaussian states. Computation of the minimal output entropy for Gaussian input is done in Section IV. Also in this section we show that the image of a fermionic product channel has a nicely ordered structure with respect to the majorization relation. The algorithm for computation of the minimal output entropy and the numerical results obtained with its help are discussed in Section V.

II. BASIC PROPERTIES OF THE FERMIONIC PRODUCT CHANNEL

A. Krauss representation

We start from constructing an explicit Krauss representation for the channel Φ defined in Eq. (4). Consider a *parity generator*

$$P = (-i)^n \hat{c}_1 \hat{c}_2 \cdots \hat{c}_{2n-1} \hat{c}_{2n}. \tag{9}$$

In terms of qubits it looks as $P = \sigma_1^z \otimes \cdots \otimes \sigma_n^z$. One can easily check that

$$\hat{c}(\mathbf{x}) P = (-1)^{|\mathbf{x}|} P \hat{c}(\mathbf{x}),$$

where $|\mathbf{x}|$ is the Hamming weight of the string $\mathbf{x}$. Introduce Krauss operators

$$K_p = iP\hat{c}_p \quad \text{such that} \quad K_p\, \hat{c}_q = (-1)^{\delta_{pq}} \hat{c}_q\, K_p.$$

A one-mode quantum channel

$$\Phi_p(\rho) = \frac{1}{2}(1+b_p)\rho + \frac{1}{2}(1-b_p) K_p\, \rho K_p$$

implements the transformation Eq. (4) with $b_q = 1$ for all $q \neq p$. Note that Φ_p is a trace preserving completely positive (TPCP) map iff $|b_p| \leq 1$. Since the one-mode channels commute, $\Phi_p \circ \Phi_q = \Phi_q \circ \Phi_p$, we conclude that

$$\Phi = \Phi_1 \circ \Phi_2 \circ \cdots \circ \Phi_{2n}. \tag{10}$$

Accordingly, Φ is a TPCP map for all $|b_p| \leq 1$. Expanding the product Eq. (10) yields the desired Krauss representation of Φ,

$$\Phi(\rho) = \sum_{\mathbf{x}} p(\mathbf{x})\, \hat{c}(\mathbf{x})\, \rho\, \hat{c}(\mathbf{x})^\dagger, \tag{11}$$

where

$$p(\mathbf{x}) = \frac{1}{2^{2n}} \prod_{q=1}^{2n} (1 + (-1)^{|\mathbf{x}|+x_q}\, b_q).$$

The channels related by a transformation $b_p \to -b_p$ for some p are unitarily equivalent, so we can focus on non-negative values $0 \leq b_p \leq 1$.

B. Covariance

The subsection summurazes some ideas proposed in the papers [8, 9]. Let $\mathcal{S} = \{U_1, U_2, \ldots, U_d\}$ be a set of n-qubit unitary operators (which may or may not constitute a group) that completely randomize any quantum state:

$$\frac{1}{d} \sum_{j=1}^{d} U_j\, \rho\, U_j^\dagger = I/2^n \quad \text{for any state} \quad \rho. \tag{12}$$

A quantum channel Θ is said to be *covariant* if it commutes with all operators $U \in \mathcal{S}$, i.e.,

$$\Theta(UAU^\dagger) = U\Theta(A)U^\dagger, \quad U \in \mathcal{S},$$

for any operator A. Covariance is a very useful property of a channel that allows one to reduce the problem of calculating the capacity to the problem of finding the minimum output entropy of a channel, see Eq. (8). Indeed, let $|\Psi\rangle$ be a state with the minimum output entropy. Consider an esemble of pure states $\{U|\Psi\rangle\}_{U\in\mathcal{S}}$, where U is chosen randomly and uniformly from the set $\mathcal{S}$. This ensemble has the maximally mixed average state since $d^{-1}\sum_{j=1}^{d} U_j|\Psi\rangle\langle\Psi|U_j^\dagger = 2^{-n} I$. Therefore the Holevo quantity of this ensemble is $n - S(\Theta(|\Psi\rangle\langle\Psi|)) = n - S_{min}(\Theta)$. Obviously, this is the absolute maximum of the Holevo quantity for any ensemble of states in $Im(\Theta)$, so that $C_1(\Theta) = n - S_{min}(\Theta)$ for any covariant channel.

The channel Φ defined in Eq. (4) is covariant. Indeed, let $\mathcal{S}$ be the set of $d = 4^n$ operators $\hat{c}(\mathbf{x})$. One can easily check that

$$\frac{1}{4^n} \sum_{\mathbf{x}} \hat{c}(\mathbf{x})\, \hat{c}(\mathbf{y})\, \hat{c}(\mathbf{x}) = \begin{cases} 0 & \text{if } \mathbf{y} \neq 0, \\ I & \text{if } \mathbf{y} = 0. \end{cases}$$

It means that $\mathcal{S}$ possesses the randomizing property Eq. (12). On the other hand it is obvious that

$$\Phi(\hat{c}(\mathbf{x}) A\, \hat{c}(\mathbf{x})^\dagger) = \hat{c}(\mathbf{x})\Phi(A)\, \hat{c}(\mathbf{x})^\dagger$$

for any operator A. Therefore Φ is covariant channel, and its one-shot capacity can be found from Eq. (7).

C. Physical motivation

Under certain circumstances the fermionic product channels may adequately describe evolution of fermi systems interacting with environment. As a toy model consider a fermi system that consists of a single fermionic mode ($n = 1$) which we describe by creation/annihilation operators $\hat{a}^\dagger$, $\hat{a}$. Suppose the system interacts with an environment which also can be described by a single fermionic mode $\hat{a}_e^\dagger$, $\hat{a}_e$. Suppose the interaction between the system and the environment is just the hopping Hamiltonian $H_{int} = g(\hat{a}^\dagger \hat{a}_e + \hat{a}_e^\dagger \hat{a})$. We assume that

the initial state of the environment is maximaly mixed, $\rho_e = (1/2)I$. The toy model channel is defined as

$$\Phi(\rho) = \mathrm{Tr}_e \left(U\rho\otimes\rho_e U^\dagger\right), \quad U = \exp\left(-iH_{int}t\right).$$

After straighforward calculations one gets

$$\Phi(I) = I, \quad \Phi(\hat{a}) = \lambda\hat{a}, \quad \Phi(\hat{a}^\dagger\hat{a} - \hat{a}\hat{a}^\dagger) = \lambda^2(\hat{a}^\dagger\hat{a} - \hat{a}\hat{a}^\dagger),$$

where $\lambda = \cos(gt)$. Introducing $\hat{c}$ generators $\hat{c}_1 = \hat{a}_1 + \hat{a}^\dagger$, $i\hat{c}_2 = \hat{a}_1 - \hat{a}_1^\dagger$ we get: $\Phi(\hat{c}_j) = \lambda\,\hat{c}_j$ and $\Phi(\hat{c}_1\hat{c}_2) = \lambda^2\,\hat{c}_1\hat{c}_2$. This is an example of the product fermionic channel Eq. (4) with $n = 1$ and $b_1 = b_2 = \lambda$. More generally, if the environment possesses a superconducting order parameter, the interaction Hamiltonian may include terms with particle-hole conversion, like in Andreev reflection of electrons. In this case one can tune the interaction to obtain an arbitrary string of coefficients b_p.

In fact, any unital channel that can be described by a quadratic interaction Hamiltonian between the system and the environment is unitarily equivalent to a channel from the family Eq. (4), see [7] for a proof.

III. GAUSSIAN STATES OF FERMIONS

In this section we describe a variational family of n-qubit states that will be used to minimize the output entropy of the channel Eq. (4). These are Gaussian states. The simplest example of a Gaussian state is the Fock vacuum $|0^{\otimes n}\rangle$ or any other product state diagonal in the standard basis. Denote

$$\rho_\lambda = \frac{1}{2^n}\prod_{j=1}^{n}(I + \lambda_j\sigma_j^z) = \frac{1}{2^n}\prod_{j=1}^{n}(I - i\lambda_j\,\hat{c}_{2j-1}\hat{c}_{2j}). \quad (13)$$

An arbitrary Gaussian state can be converted into the standard form ρ_λ by a unitary evolution with a Hamiltonian quadratic in the operators $\hat{c}_p$. Here is a strict definition.

Definition 1. *A state ρ is Gaussian iff it can be represented as*

$$\rho = U\,\rho_\lambda\,U^\dagger, \quad U = \exp\left(iH_2 + iH_1\right), \quad (14)$$

where H_2 and H_1 are Hermitian linear combinations of operators $\hat{c}_p\hat{c}_q$ and $\hat{c}_p$ respectively.

One can easily check that any pure Gaussian state has a form $U\,|0^{\otimes n}\rangle$, where U is as above.

It will be more convenient to work with even Gaussian states which correspond to choosing $H_1 = 0$ in Definition 1.

Definition 2. *A state ρ of n qubits is even Gaussian iff it can be represented as*

$$\rho = U\,\rho_\lambda\,U^\dagger, \quad U = \exp\left(iH_2\right), \quad (15)$$

where H_2 is a Hermitian linear combination of $\hat{c}_p\hat{c}_q$.

One can easily check that any pure even Gaussian state has a form $\exp(iH_2)\,|0^{\otimes n}\rangle$. Our strategy will be to prove all statements first for even Gaussian states. Then we shall show how to establish a correspondence between Gaussian states of n-qubit system and even Gaussian states of $(n+1)$-qubit system.

Unitary operators $U = \exp(iH_2)$ will be refered to as Bogolyubov tramsformations. Their conjugated action is

$$U\,\hat{c}_p\,U^\dagger = \sum_{q=1}^{2n} R_{pq}\,\hat{c}_q,$$

where R is a rotation, $R\,R^T = I$, $\det(R) = 1$. Any rotation $R \in SO(2n)$ can be realized by a proper Bogolyubov transformation U. In the rest of the section we list some basic properties of Gaussian states.

A. Wick's theorem

As in the case of Gaussian probability distributions, an even Gaussian state ρ is completely characterized by its covariance matrix,

$$M_{pq} = -\frac{i}{2}\,\mathrm{Tr}\left[\rho\left(\hat{c}_p\hat{c}_q - \hat{c}_q\hat{c}_p\right)\right].$$

All higher moments of ρ can be expressed in terms of M using Wick's theorem. Namely, for any even binary string $\mathbf{x} \in \{0,1\}^{2n}$, $|\mathbf{x}| = 2l$, one has

$$\mathrm{Tr}(\rho\,\hat{c}(\mathbf{x})) = i^l Pf(M[\mathbf{x}]), \quad (16)$$

where $M[\mathbf{x}]$ is a $2l\times 2l$ submatrix of M obtained by selecting all matrix elements M_{pq} for which $x_p = x_q = 1$, and Pf stands for the Pfaffian of a matrix. For example,

$$i^{-2}\,\mathrm{Tr}(\rho\,\hat{c}_1\hat{c}_2\hat{c}_3\hat{c}_4) = M_{12}M_{34} - M_{13}M_{24} + M_{14}M_{23}.$$

If ρ is an even Gaussian state then all odd correlators vanish,

$$\mathrm{Tr}(\rho\,\hat{c}(\mathbf{x})) = 0 \quad \text{whenever} \quad |\mathbf{x}| = 2l+1,$$

which can be easily derived from the fact that ρ commutes with the parity generator Eq. (9), $P\rho = \rho P$.

We shall use Wick's theorem to prove that the channel Φ defined in Eq. (4) maps the set of Gaussian states into itself. Also Wick's theorem will allow us to describe the action of Φ as a simple transformation of covariance matrices.

B. Admissible covariance matrices

Recall that our goal is to use Gaussian states as variational states to minimize the output entropy of the channel Φ. Since a Gaussian state is completely characterized

by its covariance matrix, we have to find a set of admissible covariance matrices. Note that if M is chosen arbitrarily, the operator ρ defined by Wick's theorem may have negative eigenvalues, i.e., it might not describe a quantum state at all.

Lemma 1. *A real antisymmetric matrix M is a covariance matrix of an even Gaussian state iff*

$$M^T M \le I. \tag{17}$$

The corresponding state is pure iff $M^T M = I$.

For a proof of the lemma see [7]. Since M is a real antisymmetric matrix, its spectrum consists of n pairs of conjugated eigenvalues $\pm i\lambda_j$, $\lambda_j \ge 0$. We shall refer to the numbers $\lambda_1, \ldots, \lambda_n$ as singular values of M. The consistency condition Eq. (17) is equivalent to inequalities $\lambda_j \le 1$. The singular values λ_j completely determine the spectrum of an even Gaussian state which is a product of n binary spectrums $(1/2)(1 \pm \lambda_j)$, see [7]. Accordingly, the von Neumann entropy $S(\rho) = -\operatorname{Tr}(\rho \log \rho)$ is

$$S(\rho) = \sum_{j=1}^{n} H\left(\frac{1+\lambda_j}{2}\right). \tag{18}$$

C. Reduction to even states

A simple correspondence between the sets of Gaussian states and even Gaussian states can be established by adding one extra fermionic mode to the system (this idea was proposed by Knill [11]). Let ρ be an n-qubit Gaussian state. Consider a linear map $\mathcal{E}$ that maps n-qubit states into $(n+1)$-qubit states according to

$$\mathcal{E}(\rho) = V\, \rho \otimes (I/2)\, V^\dagger, \quad V = \exp\left(i\frac{\pi}{4}\hat{c}_{2n+1}\right). \tag{19}$$

Here $I/2$ stands for one maximally mixed qubit which is labeled by $n+1$. Since V is a unitary operator, one has

$$S(\mathcal{E}(\rho)) = S(\rho) + 1 \quad \text{for any} \quad \rho. \tag{20}$$

Let us show that $\mathcal{E}(\rho)$ is an even Gaussian state for any Gaussian state ρ. Indeed, representing ρ in the form Eq. (14) one gets

$$\mathcal{E}(\rho) = (VU)\, \rho_\lambda \otimes (I/2)\, (VU)^\dagger,$$

where $U = \exp(iH_2 + iH_1)$. Taking into account that V commutes with ρ_λ, we can rewrite it as

$$\mathcal{E}(\rho) = (VUV^\dagger)\, \rho_\lambda \otimes (I/2)\, (VUV^\dagger)^\dagger. \tag{21}$$

Recall that H_2 is a linear combination of $\hat{c}_p\hat{c}_q$ with $1 \le p, q \le 2n$, while H_1 is a linear combination of $\hat{c}_1, \ldots, \hat{c}_{2n}$. It follows that H_2 commutes with V. On the other hand, V does not commute with H_1. Taking into account that

$$V\, \hat{c}_p\, V^\dagger = i\hat{c}_p\, \hat{c}_{2n+1} \quad \text{for any} \quad p = 1, \ldots, 2n,$$

one can easily check that $V\, H_1 V^\dagger$ is a linear combination of operators $\hat{c}_p\, \hat{c}_{2n+1}$, $p = 1, \ldots, 2n$. It means that $V(H_2 + H_1)V^\dagger$ is a linear combination of quadratic operators $\hat{c}_p\, \hat{c}_q$ only, where $1 \le p, q, \le 2n+1$. Comparing Eq. (21) and Definition 2 we conclude that $\mathcal{E}(\rho)$ is an even Gaussian state.

Conversly, if $\eta \equiv \mathcal{E}(\rho)$ is an even Gaussian state then ρ is a Gaussian state. Indeed, by definition of $\mathcal{E}$, the generator $\hat{c}_{2n+2}$ commutes with η. It means that a covariance matrix M of η has all zeroes in the column $2n+2$ and the row $2n+2$, i.e., $M_{p,2n+2} = M_{2n+2,p} = 0$. Therefore it can be represented as

$$M = R\, M_0\, R^T, \quad M_0 = \begin{pmatrix} N & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where R is a Bogolyubov transformation that involves only the generators $\hat{c}_1, \ldots, \hat{c}_{2n+1}$. Accordingly, η can be represented as

$$\eta = U\, \eta_0 \otimes (I/2)\, U^\dagger, \quad U = \exp(iH_2),$$

where η_0 is an even Gaussian state of n qubits with the covariance matrix N, while H_2 is a Hermitian linear combination of $\hat{c}_p\, \hat{c}_q$ with $1 \le p, q \le 2n+1$. Repeating the same arguments as above we get

$$\rho \otimes (I/2) = W\, \eta_0 \otimes (I/2)\, W^\dagger, \quad W = V^\dagger\, U\, V.$$

Conjugating H_2 by V we do not affect terms $\hat{c}_p\, \hat{c}_q$ that do not contain $\hat{c}_{2n+1}$. On the other hand, a term like $\hat{c}_p\hat{c}_{2n+1}$ is converted into $i\hat{c}_p$. Thus the generator $\hat{c}_{2n+1}$ does not enter into W, and we get $\rho = W\, \eta_0\, W$. It follows that ρ is a Gaussian state.

D. Action of product channels on Gaussian states

Let ρ be an even Gaussian state with a covariance matrix M and Φ be a product fermionic channel with attenuation coefficients $b_1, \ldots, b_{2n}$. Moments of the state ρ are related to moments of the state $\Phi(\rho)$ by the following identity

$$\operatorname{Tr}(\Phi(\rho)\hat{c}(\mathbf{x})) = \left(\prod_{p\,:\,x_p=1} b_p\right) \operatorname{Tr}(\rho\hat{c}(\mathbf{x})).$$

Comparing it with Wick's theorem Eq. (16) we conclude that $\Phi(\rho)$ is an even Gaussian state with a covariance matrix

$$\Phi(M) = B\, M\, B^T, \tag{22}$$

where B is a diagonal matrix with entries $b_1, \ldots, b_{2n}$,

$$B = \operatorname{diag}(b_1, b_2, \ldots, b_{2n-1}, b_{2n}).$$

Lemma 2. *Fermionic product channels map Gaussian states into Gaussian states.*

Proof. Let Φ be a channel as above. Define a fermionic product channel $\hat{\Phi}$ acting on $n+1$ qubits such that

$$\begin{aligned} \hat{\Phi}(\hat{c}_p) &= b_p\,\hat{c}_p, \quad p=1,\ldots,2n, \\ \hat{\Phi}(\hat{c}_{2n+1}) &= \hat{c}_{2n+1}, \\ \hat{\Phi}(\hat{c}_{2n+2}) &= 0. \end{aligned} \tag{23}$$

One can easily check that $\mathcal{E}$ commutes with Φ in the following sense:

$$\mathcal{E}\circ\Phi = \hat{\Phi}\circ\mathcal{E}. \tag{24}$$

Let ρ be a Gaussian state. Then $\mathcal{E}(\rho)$ is an even Gaussian state, and thus $(\hat{\Phi}\circ\mathcal{E})(\rho)$ is an even Gaussian state. It follows from Eq. (24) that $(\mathcal{E}\circ\Phi)(\rho))$ is also an even Gaussian state. As we proved above, it implies that $\Phi(\rho)$ itself is a Gaussian state. □

E. Three-qubit states

Suppose we restrict our attention only to even states, i.e., those satisfying

$$P\,|\Psi\rangle = |\Psi\rangle,$$

where P is the parity generator, see Eq. (9). Obviously, for one qubit, $n=1$, the only even state is $|0\rangle$. This is a Gaussian state. If one takes two-qubits, even states constitute a two-dimensional subspace with a basis $|0,0\rangle$ and $|1,1\rangle$. Let us consider operators $A=(-i)\hat{c}_1\hat{c}_2$, $B=(-i)\hat{c}_2\hat{c}_3$, and $C=(-i)AB=(-i)\hat{c}_3\hat{c}_1$. They all commute with P and obey the same commutation rules as the Pauli operators σ^x, σ^y, σ^z respectivly. Therefore any unitary operator on the even subspace can be represented as $U=\exp(i\alpha A+i\beta B+i\gamma C)$ for some real numbers α, β, γ. Accordingly, any even state has a form $U\,|0,0\rangle$. It follows from Definition 2 that any even two-qubit state is Gaussian. Surprisingly enough, the same is true for three qubits as well.

Lemma 3. *Any even pure state of three qubits is Gaussian.*

Proof. Let $\rho=|\Psi\rangle\langle\Psi|$ be an even pure Gaussian state of three qubits. Taking into account that all odd moments of ρ vanish and that $P\,|\Psi\rangle=|\Psi\rangle$, we can express ρ as

$$\rho=\frac{1}{8}(I+P)(I-\frac{i}{2}\sum_{p,q}M_{pq}\,\hat{c}_q\hat{c}_q),$$

where the sum runs over all $1\le p,q\le 6$ and M is a real antisymmetric matrix. It is a well-known fact from linear algebra that such a matrix M can be represented as

$$M=R\begin{pmatrix}\lambda_1\omega & 0 & 0\\ 0 & \lambda_2\omega & 0\\ 0 & 0 & \lambda_3\omega\end{pmatrix}R^T,\quad \omega=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix},$$

where λ_j are real numbers and $R\in SO(6)$ is a rotation. Consider a Bogolyubov transformation U such that $U\,\hat{c}_p\,U^\dagger=\sum_{q=1}^{8}R_{pq}\,\hat{c}_q$. Then

$$\rho'=U\,\rho\,U^\dagger=\frac{1}{8}(I+P)(I-i\lambda_1\,\hat{c}_1\hat{c}_2-i\lambda_2\,\hat{c}_3\hat{c}_4-i\lambda_3\,\hat{c}_5\hat{c}_6).$$

Here we used the fact that $UPU^\dagger=\det(R)P=P$. The state ρ' is diagonal in the standard basis. On the other hand ρ' is a pure state. Therefore ρ' is a vector of the standard basis with even number of ones. One can easily check that any such vector is a Gaussian state. It follows that ρ is a Gaussian state as well. □

The minimum number of qubits supporting even pure non-Gaussian states is $n=4$. For example, consider a state

$$|\theta\rangle=\cos(\theta)\,|0,0,0,0\rangle+\sin(\theta)\,|1,1,1,1\rangle.$$

One can easily check that the covariance matrix of $|\theta\rangle$ has singular values smaller than one whenever $\sin(2\theta)\neq 0$. It follows from Lemma 1 that such states are not Gaussian.

IV. COMPUTATION OF THE GAUSSIAN CAPACITY

Throught this section Φ is a fermionic product channel with attenuation coefficients $b_1,\ldots,b_{2n}$ ordered in non-increasing way, i.e.,

$$b_1\ge b_2\ge\ldots\ge b_{2n}.$$

A computation of the Gaussian capacity will proceed in three steps. Firstly we shall compute the minimum output entropy achievable on even Gaussian input states. Then we shall generalize it to arbitrary Gaussian input using the trick with addition of a qubit. Finally we shall verify that the optimal signaling ensemble consists of Gaussian states.

A. Even Gaussian input

Our intermediate goal is to find the minimal value of the output von Neumann entropy $S(\Phi(\rho))$ provided that the input ρ is an even Gaussian state. Denote this minimal value $S_{min,e}(\Phi)$.

Let M be the covariance matrix of ρ. Since the minimum output entropy is achieved on a pure input state, we can assume that

$$M^T\,M=I, \tag{25}$$

see Lemma 1. As we have shown in Section III, the output state $\Phi(\rho)$ is an even Gaussian state with a covariance matrix

$$\Phi(M)=B\,M\,B^T,\quad B=\mathrm{diag}(b_1,\ldots,b_{2n}). \tag{26}$$

Let $\lambda_1, \ldots, \lambda_n$ be singular values of the matrix $\Phi(M)$ (recall that its eigenvalues are $\pm i\lambda_j$). Taking into account Eq. (18) we get

$$S_{min,e} = \min_{\lambda_1,\ldots,\lambda_n} \sum_{j=1}^{n} H\left(\frac{1+\lambda_j}{2}\right),$$

where the minimum is taken over all strings of singular values $\lambda_1, \ldots, \lambda_n$ that are consistent with Eqs. (25,26).

We claim that

$$S_{min,e}(\Phi) = \sum_{j=1}^{n} H\left(\frac{1+b_{2j-1}b_{2j}}{2}\right). \qquad (27)$$

The proof will be based on convexity arguments and two basic facts from the majorization theory, see [12]. Recall that if $y = (y_1, \ldots, y_n)$ and $z = (z_1, \ldots, z_n)$ are strings of real numbers, the majorization relation $y \prec z$ is equivalent to inequalities

$$\sum_{j=1}^{k} y_j^{\downarrow} \leq \sum_{j=1}^{k} z_j^{\downarrow}, \quad j = 1, \ldots, n \qquad (28)$$

with an equality for $k = n$. Here $y_j^{\downarrow}$ and $z_j^{\downarrow}$ are the j-largest elements of the strings y and z respectively.

Fact 1: If $f(x)$ is a concave function of a real variable x and $(y_1, \ldots, y_n) \prec (z_1, \ldots, z_n)$ then

$$\sum_{j=1}^{n} f(y_j) \geq \sum_{j=1}^{n} f(z_j).$$

Fact 2: (Horn-Visser-Zaanen theorem)
For any square matrix A define $D_l(A)$ as a product of l largest singular values of A. Then for any $m \times m$ matrices A and B

$$D_l(AB) \leq D_l(A)\, D_l(B) \quad \text{for all} \quad l = 1, \ldots, m. \qquad (29)$$

Note that $D_m(A) = |\det(A)|$, so Eq. (29) becomes an equality for $l = m$.

Now let us proceed to a proof of Eq. (27). Let $\lambda_j^{\downarrow}$ be the sequence of singular values λ_j ordered in non-increasing way. Taking into account that

$$D_l(B) = D_l(B^T) = \prod_{j=1}^{l} b_j, \quad D_l(M) = 1,$$

we can rewrite the Horn-Visser-Zaanen inequality

$$D_{2k}(B\, M\, B^T) \leq D_{2k}(B)\, D_{2k}(M)\, D_{2k}(B^T)$$

as

$$\prod_{j=1}^{k} (\lambda_j^{\downarrow}) \leq \prod_{p=1}^{2k} b_p, \qquad (30)$$

with equality for $k = n$. Let us introduce auxiliary variables $y_j = \log(\lambda_j^{\downarrow})$ and $z_j = \log(b_{2j-1}b_{2j})$ taking values on the interval $(-\infty, 0]$. Obviously, $y_1 \geq \ldots \geq y_n$ and $z_1 \geq \ldots \geq z_n$. The inequalities Eq. (30) become equivalent to the majorization inequalities Eq. (28), so that

$$(y_1, \ldots, y_n) \prec (z_1, \ldots, z_n).$$

The von Neumann entropy of the output state is

$$S(\Phi(\rho)) = \sum_{j=1}^{n} f(y_j), \quad f(x) = H\left(\frac{1-e^x}{2}\right).$$

To prove that the function $f(x)$ is concave let us represent it as $f(x) = H(g(x))$, where $g(x) = (1/2)(1-e^x)$. Obviously, $g(x)$ is concave and maps the interval $(-\infty, 0]$ into the interval $[0, 1/2)$. The function $H(x)$ is concave and monotone increasing on the interval $[0, 1/2]$. Therefore $H(g(x))$ is concave on the interval $(-\infty, 0]$. Applying Fact 1 we get

$$\begin{aligned} S(\Phi(\rho)) &= \sum_{j=1}^{n} f(y_j) \geq \sum_{j=1}^{n} f(z_j) \\ &= \sum_{j=1}^{n} H\left(\frac{1+b_{2j-1}b_{2j}}{2}\right). \end{aligned} \qquad (31)$$

This value of the output entropy can be achieved if the input ρ is the Fock vacuum, i.e.,

$$\rho = \frac{1}{2^n} \prod_{j=1}^{n} (I - i\hat{c}_{2j-1}\hat{c}_{2j}) = |0^{\otimes n}\rangle\langle 0^{\otimes n}|.$$

Indeed, the corresponding output state is

$$\begin{aligned} \Phi(\rho) &= \frac{1}{2^n} \prod_{j=1}^{n} (I - ib_{2j-1}b_{2j}\hat{c}_{2j-1}\hat{c}_{2j}) \\ &= \frac{1}{2^n} \prod_{j=1}^{n} (I + b_{2j-1}b_{2j}\sigma_j^z). \end{aligned} \qquad (32)$$

$$(33)$$

The von Neumann entropy of $\Phi(\rho)$ matches the lower bound Eq. (31).

B. Proof of Theorem 1

Now we are ready to prove Theorem 1. Firstly we shall find the minimum output entropy achievable on a Gaussian input (not necesserily even). Denote this quantity $S_{min,g}(\Phi)$. We claim that

$$S_{min,g}(\Phi) = H\left(\frac{1+b_1}{2}\right) + \sum_{j=1}^{n-1} H\left(\frac{1+b_{2j}b_{2j+1}}{2}\right). \qquad (34)$$

Indeed, let G_n and G_n^e be sets of n-qubit Gaussian states and even Gaussian states respectively, such that

$$S_{min,g} = \min_{\rho \in G_n} S(\Phi(\rho)).$$

We shall make use of the linear map $\mathcal{E}$ defined in Eq. (19) that adds one qubit to the system. Taking into account Eq. (20) we get

$$S_{min,g} = \min_{\rho\in G_n} S((\mathcal{E}\circ\Phi)(\rho)) - 1.$$

Applying the commutation rule Eq. (24) we arrive to

$$S_{min,g} = \min_{\rho\in G_n} S((\hat{\Phi}\circ\mathcal{E})(\rho)) - 1.$$

As we have shown in Section III.C, $\mathcal{E}(\rho)$ is an even Gaussian state, so that

$$S_{min,g} \geq \min_{\eta\in G^e_{n+1}} S(\hat{\Phi}(\eta)) - 1.$$

Since we already now how to compute the minimum over even Gaussian states, see Eq. (27), one gets

$$S_{min,g} \geq \sum_{j=1}^{n+1} H\left(\frac{1+\hat{b}^{\downarrow}_{2j-1}\hat{b}^{\downarrow}_{2j}}{2}\right) - 1.$$

Here $\hat{b}_1,\ldots,\hat{b}_{2n+2}$ are the attenuation coefficients for the map $\hat{\Phi}$, see Eq. (23). Obviosuly, the largest of them is $\hat{b}^{\downarrow}_1 = \hat{b}_{2n+1} = 1$, while the smallest one is $\hat{b}^{\downarrow}_{2n+2} = \hat{b}_{2n+2} = 0$. Therefore we arrive to

$$S_{min,g} \geq H\left(\frac{1+b_1}{2}\right) + \sum_{j=1}^{n-1} H\left(\frac{1+b_{2j}\,b_{2j+1}}{2}\right). \quad (35)$$

This lower bound is achieved on a state

$$\rho_* = \frac{1}{2^n}(I+\hat{c}_1)(I-i\hat{c}_2\hat{c}_3)\cdots(I-i\hat{c}_{2n-2}\hat{c}_{2n-1}). \quad (36)$$

We have to verify that ρ_* is indeed a Gaussian state in the sense of Definition 1. Indeed, one can easily check that

$$\rho_* = (V_1V_2)\,|0^{\otimes n}\rangle\langle 0^{\otimes n}|\,(V_1V_2)^{\dagger},$$

where V_2 is a Bogolyubov transformation that shifts the operators $\hat{c}_p$ cyclically according to

$$(\hat{c}_1,\hat{c}_2,\hat{c}_3,\ldots,\hat{c}_{2n}) \to (-\hat{c}_{2n},\hat{c}_1,\hat{c}_2,\ldots,\hat{c}_{2n-1}),$$

while $V_1 = \exp(-i\frac{\pi}{4}\hat{c}_{2n})$. The operator V_1 is chosen such that $V_1(\hat{c}_{2n}\hat{c}_1)V_1^{\dagger} = -i\hat{c}_1$. One remains to observe that

$$\Phi(\rho_*) = \frac{1}{2^n}(I+b_1\hat{c}_1)\prod_{j=1}^{n-1}(I-ib_{2j}b_{2j+1}\hat{c}_{2j}\hat{c}_{2j+1}).$$

All factors in this product can be diagonalized simultaniously, so the spectrum of $\Phi(\rho_*)$ is a product of n binary spectrums $(1/2)(1\pm b_1)$ and $(1/2)(1\pm b_{2j}b_{2j+1})$. The corresponding entropy $S(\Phi(\rho_*))$ matches the lower bound Eq. (35).

It is obvious that $C_1^g \leq n - S_{min,g}$. The optimal signaling ensemble that achieves this upper bound consists of 2^n Gaussian states

$$\frac{1}{2^n}(I\pm\hat{c}_1)(I\pm i\hat{c}_2\hat{c}_3)\cdots(I\pm i\hat{c}_{2n-2}\hat{c}_{2n-1}).$$

These states constitute an orthonormal basis and each of them yields the output entropy $S_{min,g}$.

To complete the proof of Theorem 1 one remains to show that $C_1 = C_1^g$ for $n = 2$, or, equivalently, that $S_{min} = S_{min,g}$ for two-qubit channels. Denote D_n and D_n^e sets of all n-qubits states and even states respectively. Repeating the same arguments as above one gets

$$\min_{\rho\in D_2} S(\Phi(\rho)) \geq \min_{\eta\in D_3^e} S(\hat{\Phi}(\eta)) - 1.$$

Obviously, the minimum in the righthand side is achieved on a pure state η. But we have already shown in Section III.D that any pure even three-qubit state is Gaussian. Therefore one can substitute the minimum over D_3^e by a minimum over G_3^e and proceed as above. We have proved Theorem 1.

C. Minimal entropy and majorization

In the previous subsection we have found a Gaussian state ρ_*, see Eq. (36), such that $S(\Phi(\rho_*)) \leq S(\Phi(\rho))$ for any other Gaussian state ρ. Since the von Neumann entropy measures amount of randomness contained in a state, we can expect that $\Phi(\rho^*)$ is the least randomized state in the image of Φ (as far as Gaussian inputs are concerned). Here we shall put this statement into a more strict form and prove the following interesting fact.

Proposition 1:

$$\Phi(\rho) \prec \Phi(\rho_*) \quad \text{for any Gaussian } \rho. \quad (37)$$

Here the majorization relation between the density operators $\rho \prec \eta$ means that the spectrum of ρ is majorized by the spectrum of η. The majorization relation $\rho \prec \eta$ is the quantative version of the statement "ρ is more randomized than η". It is well known that $\rho \prec \eta$ iff $\rho = \sum_\alpha p_\alpha U_\alpha \eta U_\alpha^{\dagger}$ for some unitaries U_α and some probability distribution p_α. In particular $\rho \prec \eta$ implies $S_\alpha(\rho) \geq S_\alpha(\eta)$, where $S_\alpha(\rho) = \frac{1}{1-\alpha}\log(\mathrm{Tr}\,\rho^\alpha)$ is the Rényi entropy.

Proposition 1 is a simple consequence of the inequality Eq. (30) and the following technical lemma.

Lemma 4. *Let $\alpha_1,\ldots,\alpha_n$ and $\beta_1,\ldots,\beta_n$ be real numbers satisfying inequalities*

$$1 \geq \alpha_1 \geq \ldots \geq \alpha_n \geq 0, \quad 1 \geq \beta_1 \geq \ldots \geq \beta_n \geq 0,$$

Let $x = (x_1,\ldots,x_n)$ and $y = (y_1,\ldots,y_n)$ be random n-bit variables with probability distributions

$$P(x) = \frac{1}{2^n}\prod_{j=1}^{n}(1+(-1)^{x_j}\alpha_j),$$

and

$$Q(y) = \frac{1}{2^n} \prod_{j=1}^{n} (1 + (-1)^{y_j} \beta_j).$$

If for all $k = 1, \ldots, n$ *one has*

$$\prod_{j=1}^{k} \alpha_j \leq \prod_{j=1}^{k} \beta_j$$

with an equality for $k = n$ *then* P *is majorized by* Q*,*

$$P \prec Q.$$

Let us first show how to derive Eq. (37) from the lemma.

Proof of the proposition. It will be convenient to use a notation $P_n(\alpha_1, \ldots, \alpha_n)$ for a product of n binary distributions $(1/2)(1 \pm \alpha_j)$. One suffices to prove Eq. (37) for pure states ρ. As we have shown in the subsection A, $\Phi(\rho)$ is an even Gaussian state whose covariance matrix has singular values $\lambda_1^{\downarrow}, \ldots, \lambda_n^{\downarrow}$ obeying the inequality Eq. (30). As was mentioned in Section III.B, the spectrum of $\Phi(\rho)$ is a product of n binary spectrums $(1/2)(1 \pm \lambda_j^{\downarrow})$. Applying Lemma 4 to $\alpha_j = \lambda_j^{\downarrow}$ and $\beta_j = b_{2j-1} b_{2j}$ we conclude that

$$\Phi(\rho) \prec P_n(b_1 b_2, \ldots, b_{2n-1} b_{2n}). \tag{38}$$

Now we use the correspondence between Gaussian n-qubit states and even Gaussian $(n+1)$-qubit states. Let ρ be an arbitrary Gaussian state. Consider a linear map $\mathcal{E}$ defined in Eq. (19) and a channel $\hat{\Phi}$ defined in Eq. (23). Recall that $\mathcal{E} \circ \Phi = \hat{\Phi} \circ \mathcal{E}$. A state

$$\eta = (\mathcal{E} \circ \Phi)(\rho) = (\hat{\Phi} \circ \mathcal{E})(\rho)$$

is even Gaussian. Applying Eq. (38) to the channel $\hat{\Phi}$ we get

$$Spec(\eta) \prec P_{n+1}(b_1, b_2 b_3, \ldots, b_{2n-2} b_{2n-1}, 0). \tag{39}$$

Here $Spec(\eta)$ is the string of eigenvalues of η. On the other hand, the spectrum of η and the spectrum of $\Phi(\rho)$ are related by

$$Spec(\eta) = Spec(\Phi(\rho)) \times \{\frac{1}{2}, \frac{1}{2}\},$$

since $\mathcal{E}$ just adds one maximally mixed qubit to $\Phi(\rho)$ and then applies a unitary operator that does not change the spectrum. Taking into account Eq. (39) we get

$$\begin{aligned} Spec(\Phi(\rho)) \times \{\frac{1}{2}, \frac{1}{2}\} \prec \\ P_{n+1}(b_1, b_2 b_3, \ldots, b_{2n-2} b_{2n-1}, 0) \\ = P_n(b_1, b_2 b_3, \ldots, b_{2n-2} b_{2n-1}) \times \{\frac{1}{2}, \frac{1}{2}\}. \end{aligned} \tag{40}$$

This majorization relation involves 2^{n+1} inequalities. If one picks out only 2^n *even* inequalities (i.e. those including a sum of two, four, six, e.t.c. largest eigenvalues) one gets exactly the majorization relation

$$Spec(\Phi(\rho)) \prec P_n(b_1, b_2 b_3, \ldots, b_{2n-2} b_{2n-1}).$$

It is equivalent to the statement of the proposition. □

Proof of Lemma 4. We shall first prove the lemma for $n = 2$ by brute force method. Then we shall use standard majorization theorems to extend the proof to all $n > 2$.

If $n = 2$ then the ordered probability distributions $P = \{p_1^{\downarrow}, \ldots, p_4^{\downarrow}\}$ and $Q = \{q_1^{\downarrow}, \ldots, q_4^{\downarrow}\}$ are

$$\begin{array}{rclrcl} p_1^{\downarrow} &=& \frac{1}{4}(1+\alpha_1)(1+\alpha_2), & q_1^{\downarrow} &=& \frac{1}{4}(1+\beta_1)(1+\beta_2), \\ p_2^{\downarrow} &=& \frac{1}{4}(1+\alpha_1)(1-\alpha_2), & q_2^{\downarrow} &=& \frac{1}{4}(1+\beta_1)(1-\beta_2), \\ p_3^{\downarrow} &=& \frac{1}{4}(1-\alpha_1)(1+\alpha_2), & q_3^{\downarrow} &=& \frac{1}{4}(1-\beta_1)(1+\beta_2), \\ p_4^{\downarrow} &=& \frac{1}{4}(1-\alpha_1)(1-\alpha_2), & q_4^{\downarrow} &=& \frac{1}{4}(1-\beta_1)(1-\beta_2). \end{array}$$

The condition of the lemma for $n = 2$ looks as

$$\alpha_1 \leq \beta_1, \quad \alpha_1 \alpha_2 = \beta_1 \beta_2.$$

Our goal is to prove that $P \prec Q$. After a simple algebra the first majorization inequality $p_1^{\downarrow} \leq q_1^{\downarrow}$ can be written as

$$\alpha_1 + \frac{C}{\alpha_1} \leq \beta_1 + \frac{C}{\beta_1}, \quad C \equiv \alpha_1 \alpha_2 = \beta_1 \beta_2. \tag{41}$$

It is satisfied because a function $f(t) = t + C/t$ is monotone increasing on the interval $t \geq \sqrt{C}$ and because $\alpha_1, \beta_1 \geq \sqrt{C}$. The second majorization inequality $p_1^{\downarrow} + p_2^{\downarrow} \leq q_1^{\downarrow} + q_2^{\downarrow}$ is equivalent to $\alpha_1 \leq \beta_1$. The third majorization inequality $p_1^{\downarrow} + p_2^{\downarrow} + p_3^{\downarrow} \leq q_1^{\downarrow} + q_2^{\downarrow} + q_3^{\downarrow}$ reduces to Eq. (41) that we have already proved. Thus $P \prec Q$.

To prove the lemma for arbitrary $n > 2$ we shall need two more basic facts from the majorization theory [12]:

Definition 3. *Let* $y = (y_1, \ldots, y_n)$ *and* $z = (z_1, \ldots, z_n)$ *be strings of real numbers. A string* y *is a* T*-transform of string* z *iff there exist a real number* $0 \leq \lambda \leq 1$ *and integers* $1 \leq a < b \leq n$ *such that*

$$y_c = z_c \quad \textit{for all} \quad c \neq a, b,$$

and

$$\begin{bmatrix} y_a \\ y_b \end{bmatrix} = \begin{bmatrix} \lambda & 1-\lambda \\ 1-\lambda & \lambda \end{bmatrix} \cdot \begin{bmatrix} z_a \\ z_b \end{bmatrix}.$$

Fact 3: $y \prec z$ iff z can be converted to y by a sequence of T-transforms.

Fact 4: $y \prec z$ iff $y_j = \sum_{k=1}^{n} P_{jk} z_k$ for some doubly stochastic matrix P.

Consider auxiliary variables $y_j = \log(\alpha_j)$ and $z_j = \log(\beta_j)$. The condition of the lemma is equivalent to $y \prec z$. Keeping in mind Fact 3, one suffices to prove that $P \prec Q$ whenever a string $y = (y_1, \ldots, y_n)$ is a T-transform of a string $z = (z_1, \ldots, z_n)$. Let a, b, λ be as in

Definition 3. Since a T-transform is a doubly-stochastic matrix, one has

$$(y_a, y_b) \prec (z_a, z_b).$$

In terms of the original variables α and β it means that

$$\max(\alpha_a, \alpha_b) \leq \max(\beta_a, \beta_b), \quad \alpha_a \alpha_b = \beta_a \beta_b.$$

Applying the $n = 2$ proof given above to the marginal probability distributions P_{ab} and Q_{ab} of the bits a, b we conclude that $P_{ab} \prec Q_{ab}$. Besides, for any bit $c \neq a, b$ the marginal distributions P_c and Q_c coincide. Fact 3 implies that Q_{ab} can be converted to P_{ab} by a doubly stochastic matrix. Therefore Q can be converted to P by a doubly stochastic matrix and thus $P \prec Q$. □

V. NUMERICAL SIMULATIONS

As we have seen before, the smallest number of qubits for which the Gaussian analysis fails is $n = 3$. Therefore, we shall carry out numerical simulations for fermionic product channels acting on three and four qubits. Our goal is to verify whether the minimum of the output entropy is achieved on a Gaussian input state. Below we describe an algorithm that in many cases allows one to find the minimum output entropy with a high precision.

A. Iterative minimization of the output entropy

We shall describe an algorithm that generates a sequence of states $|\psi_k\rangle$, $k = 0, 1, 2, \ldots$, such that

- The output entropies $S_k = S(\Phi(|\psi_{k+1}\rangle\langle\psi_{k+1}|))$ are non-increasing, i.e. $S_{k+1} \leq S_k$.
- The algorithm converges to an extremal point of a functional $S(\Phi(|\psi\rangle\langle\psi|))$.

The algorithm exploits the following simple observation. Consider a functional

$$h(\rho, \eta) = -\operatorname{Tr}[\Phi(\rho)\, \log \Phi(\eta)]$$

that depends upon a pair of density operators ρ and η. Let us show that the global minimum of h coincides with S_{min}. Indeed, let us use an identity

$$h(\rho, \eta) = S(\Phi(\rho)) + S(\Phi(\rho)||\Phi(\eta)), \tag{42}$$

where $S(\omega||\sigma) = \operatorname{Tr}[\omega(\log \omega - \log \sigma)]$ is the relative entropy. Since the relative entropy is non-negative, $S(\omega||\sigma) \geq 0$ for any ω, σ, and $S(\omega||\sigma) = 0$ iff $\omega = \sigma$, we conclude that

$$S_{min}(\Phi) = \min_{\rho,\eta} h(\rho, \eta). \tag{43}$$

The key observation is that finding the minimum of $h(\rho, \eta)$ with respect to each *individual* variable is an exactly solvable problem. Indeed, if one fixes η then

$$\begin{aligned} \min_{\rho} h(\rho, \eta) &= -\max_{\rho} \operatorname{Tr}[\rho\, \Phi^*(\log \Phi(\eta))] \\ &= -\lambda_{max}(\Phi^*(\log \Phi(\eta))), \end{aligned} \tag{44}$$

where Φ^* is the linear map conjugated to Φ (with respect to the trace inner product) and $\lambda_{max}(X)$ is the maximum eigenvalue of an operator X. The optimal state ρ is obviously the highest eigenvector of $\Phi^*(\log \Phi(\eta))$. On the other hand, if one fixes ρ and minimizes $h(\rho, \eta)$ over η, the absolute minimum is achieved at $\rho = \eta$, as immediately follows from Eq. (42).

We shall attempt to find the minimum of $h(\rho, \eta)$ by carring out a sequence of alternating one-variable minimizations. This iterative minimization procedure generates a sequence of pure states $|\psi_0\rangle, |\psi_1\rangle, \ldots$, such that (i) the initial state $|\psi_0\rangle$ is chosen randomly; (ii) $|\psi_{k+1}\rangle$ is the highest eigenvector of an operator $\Phi^*(\log \Phi(|\psi_k\rangle\langle\psi_k|))$ (if the highest eigenvalue is degenerate, $|\psi_{k+1}\rangle$ is chosen randomly from the corresponding eigenspace). Consider a sequence

$$S_k = S(\Phi(|\psi_k\rangle\langle\psi_k|)), \quad k = 1, 2, \ldots, \infty.$$

Since each one-variable minimization does not increase the objective function, we have $S_{k+1} \leq S_k$, i.e., the sequence $\{S_k\}$ is monotone decreasing.

Denote $S_* = \lim_{k\to\infty} S_k$. Let $|\psi_*\rangle$ be any limiting point of the sequence $\{|\psi_k\rangle\}$ and $\rho_* = |\psi_*\rangle\langle\psi_*|$. Obviously

$$h(\rho_*, \rho_*) \leq h(\rho, \rho_*), \quad \text{and} \quad h(\rho_*, \rho_*) \leq h(\rho_*, \eta)$$

for any states ρ and η. Therefore a point $\rho = \eta = \rho_*$ is an extremum of $h(\rho, \eta)$. Thus ρ_* is an extremal point of the output entropy $S(\Phi(\rho)) = h(\rho, \rho)$. Obviously, ρ_* can not be a (local) maximum of h. However, it might be a local minimum or a saddle point. In principle, saddle points can be eliminated by computing a Hessian of $h(\rho, \rho)$ at the point ρ_*.

It seems very unlikely that the iterations can be trapped in a very narrow and shallow local minimum, since one-variable minimization allows one to 'tunnel' through potential barriers. We can also expect that the method should work well for channels that are sufficiently noisy, since a landscape of the objective function $S(\Phi(|\psi\rangle\langle\psi|))$ becomes more flat as Φ becomes less noisy. Symmetry of a channel also might be an important issue because one-variable iterative minimization typically works well for highly symmetric functions.

Before applying the algorithm to fermionic product channels we have tested it on products of one-qubit unital channels for which the exact value of S_{min} is known, see (?). The algorithm always converges to the correct value of S_{min} for channels with sufficiently large amount of noise. For example, let Θ be a one-qubit Pauli depolarizing channel, such that errors σ^x, σ^y, and σ^z occur

with a probability p each. Then the iterative algorithm correctly computes $S_{min}(\Theta^{\otimes 4}) = 4H(1-2p)$ in a region $p \geq 0.05$.

B. Numerical simulations

In the rest of this section we discuss results of numerical simulations performed for several families of fermionic product channels on three and four qubits. For each particular channel the minimum output entropy has been evaluated using the iterative minimization algorithm with 64 iterations and 16 independent choices of the initial state.

A plot on Figure 1 consists of two groups of points marked by '+' and '×'. Each of these point corresponds to some particular 3-qubit channel from the family Eq. (4). The points '+' correspond to a symmetric channel

$$\Phi_+(\hat{c}_p) = b\,\hat{c}_p, \quad p = 1, \ldots, 6, \tag{45}$$

where $0 \leq b \leq 1$ is a parameter specifying the channel. The points '×' correspond to a channel with a non-trivial distribution of the coefficients b_p, namely

$$\Phi_\times(\hat{c}_p) = b^{p/3}, \quad p = 1, \ldots, 6. \tag{46}$$

The parameter b is plotted on the horizontal axis. On the vertical axis we plotted the smallest output entropy S_{min} found by the algorithm. The dotted lines represent the minimum output entropy that can be achieved on a Gaussian input state, i.e., the function $S_{min,g}$ given by Eq. (34). The inset plot shows the deviation $S_{min} - S_{min,g}$.

It turns out that the algorithm always converges to the same value of S_{min} for the symmetric channel Φ_+. A typical deviation $S_{min} - S_{min,g}$ for this channel is about 10^{-9} which has the same order of magnitude as the numerical noise.

As for the non-symmetric channel $\Phi_\times$, the algorithm always converges to the same value of S_{min} unless b is close to 1. A typical deviation $S_{min} - S_{min,g}$ is about 10^{-9} in the region $0 \leq b \leq 0.7$, so it is not shown on the plot. The deviation is always positive however. On the other hand, if b is close to 1, the iterations sometimes converge to a local minimum with a value of S_{min} exceeding $S_{min,g}$ by several percents.

Numerical simulations have been also carried out for analogous 4-qubit channels. A plot on Figure 2 shows the smallest output entropy found by the algorithm for a symmetric channel Φ_+ defined as

$$\Phi_+(\hat{c}_p) = b\,\hat{c}_p, \quad p = 1, \ldots, 8, \tag{47}$$

and for a channel $\Phi_\times$ with a non-trivial distribution of the coefficients b_p, namely

$$\Phi_\times(\hat{c}_p) = b^{p/4}, \quad p = 1, \ldots, 8. \tag{48}$$

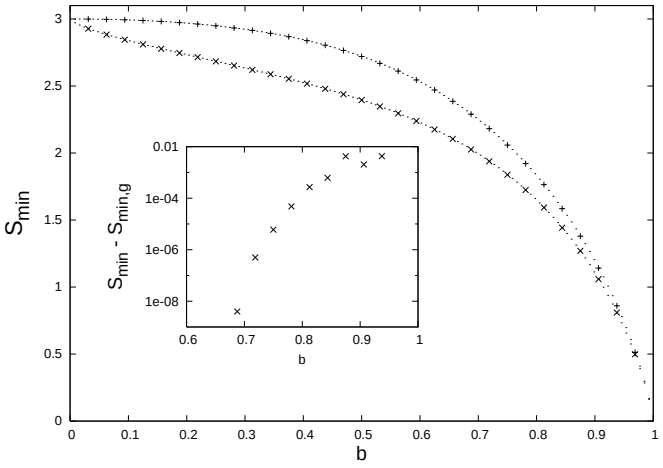

FIG. 1: Three-qubit channels. Vertical axis: the smallest output entropy S_{min} found by the iterative minimization algorithm for the channels Φ_+ and $\Phi_\times$. Horizontal axis: the parameter b specifying the channels, see Eqs. (45,46). Dotted lines: the minimum output entropy achieved on a Gaussian input ($S_{min,g}$). On the inset plot: deviation $S_{min} - S_{min,g}$ for the channel $\Phi_\times$.

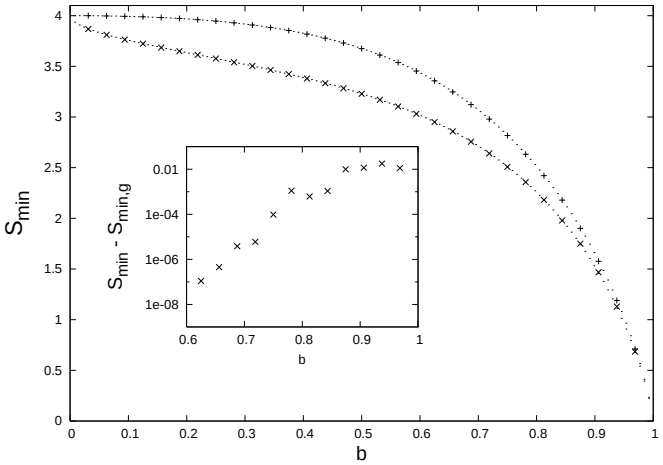

FIG. 2: Four-qubit channels. Vertical axis: the smallest output entropy S_{min} found by the iterative minimization algorithm for the channels Φ_+ and $\Phi_\times$. Horizontal axis: the parameter b specifying the channels, see Eqs. (47,48). Dotted lines: the minimum output entropy achieved on a Gaussian input ($S_{min,g}$). On the inset plot: deviation $S_{min} - S_{min,g}$ for the channel $\Phi_\times$.

Here $0 \leq b \leq 1$ is a parameter specifying the channel.

The algorithm always converges to the same value of S_{min} for the symmetric channel Φ_+. A typical deviation $S_{min} - S_{min,g}$ for this channel is about 10^{-9}.

For the channel $\Phi_\times$ the algorithm always converges to the same value of S_{min} in a region $b \leq 0.55$. The deviation $S_{min} - S_{min,g}$ drops down to 10^{-9} in a region $b \leq 0.6$, so it is not shown on the plot.

The numerical results seem to support the conjecture that the minimum output entropy is achieved on a Gaus-

sian input state, i.e., that $S_{min}(\Phi) = S_{min,g}(\Phi)$ for all fermionic product channels. The fact that the algorithm always converges to the same value of S_{min} for the symmetric channels Φ_+ suggests that a landscape of the corresponding objective function $S(\Phi_+(|\psi\rangle\langle\psi|))$ is particularly simple. It might be an indication that S_{min} can be computed analytically for the symmetric channels.

VI. ACKNOWLEDGMENTS

Discussions with Christopher King and Frank Verstraete are gratefully acknowledged. This work was supported by the National Science Foundation under grant number EIA-0086038.

[1] B. Schumacher and M. Westmoreland, *Phys. Rev.* A 56, no. 1, p. 131 (1997).
[2] A. S. Holevo, *IEEE Transactions on Information Theory* 44(1), p. 269 (1998).
[3] C. King, *J. Math. Phys.* 43, no 10, p. 4621 (2002).
[4] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J.H. Shapiro, and H.P. Yuen, *Phys. Rev. Lett.* 92, 027902 (2004).
[5] A.S. Holevo and R.F. Werner, *Phys. Rev.* A 63, 032312 (2001).
[6] B. Terhal and D. DiVincenzo, *Phys. Rev.* A 65, 032325 (2002).
[7] S. Bravyi, *Quantum Inf. and Comp.* 5, no. 3, p. 216 (2005).
[8] A.S. Holevo, quant-ph/0212025.
[9] J. Cortese, quant-ph/0211093.
[10] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, and J.H. Shapiro, *Phys. Rev.* A 70, 032315 (2004).
[11] E. Knill, quant-ph/0108033.
[12] A. Marshall and I. Olkin, "Inequalities: Theory of Majorization and Its Applications", Academic Press (1979).